R E P O R T R E S U M E S

ED 020 720 JC 680 223

ANOTHER LOOK AT TEACHER EVALUATION IN AMERICAN JUNIOR COLLEGES.

BY- KILPATRICK, GORDON

EL CAMINO COLL., EL CAMINO COLLEGE, CALIF.

REPORT NUMBER ECC-67-8 PUB DATE MAR 67

EDRS PRICE MF-$0.25 HC-$1.60 38P.

DESCRIPTORS- *JUNIOR COLLEGES, *TEACHER EVALUATION, *EFFECTIVE TEACHING, *CRITICAL INCIDENTS METHOD, TEACHER IMPROVEMENT, EVALUATION, OBSERVATION,

THE PURPOSES OF EVALUATION ARE TO (1) PERMIT THE SCHOOL TO ASSESS ITS EFFECTIVENESS, (2) IMPROVE INSTRUCTION, AND (3) ASSURE GROWTH AND DEVELOPMENT OF THE TEACHER. IF EVALUATION CAN BE PROPOSED AS A WAY TO RECOGNIZE SUPERIOR TEACHERS RATHER THAN AS A SEARCH FOR FAULT, IT IS LIKELY TO MEET WITH LESS RESISTANCE. INITIAL EVALUATION, A PART OF THE HIRING PROCESS, IS MADE FROM TRANSCRIPTS, RECOMMENDATIONS, PLACEMENT FILES, APPLICATION FORMS, AND INTERVIEWS. SUBSEQUENT EVALUATION MUST BE MADE OF THE INSTRUCTOR'S TEACHING EFFECTIVENESS. IT MAY DISCOVER NEW IDEAS OR TECHNIQUES TO BE SHARED WITH OTHERS AS WELL AS FAULTS TO BE CORRECTED. IT CAN BE ACCOMPLISHED IN WAYS OF VARYING DIFFICULTY--BY CHECKING THE LATER SCHOOL WORK OR EMPLOYMENT RECORD OF THE STUDENT, BY CORRELATING OVER THE YEARS THE STUDENT'S REALIZED SUCCESS WITH THE TEACHER'S PREDICTION OF HIS SUCCESS, CONSISTENT USE OF STANDARDIZED TESTS IN THE TEACHER'S SUBJECT AREA, BY CLASSROOM VISITS (TO OBSERVE THE PRESENTATION OF THE MATERIAL AND ITS ADAPTATION TO THE READINESS OF THE CLASS), OR BY USE OF THE "CRITICAL INCIDENT TECHNIQUE," WHICH IS THE ANALYSIS OF INCIDENTS WHERE THE TEACHER WAS JUDGED UNUSUALLY EFFECTIVE OR INEPT. EVALUATION IS WARRANTED PARTICULARLY BEFORE CONSIDERATION FOR PROMOTION, BEFORE AWARD OF TENURE, OR AFTER AN ANALYSIS OF CRITICAL INCIDENTS. (HH)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

EL CAMINO COLLEGE
ECC 67-8
RESEARCH OFFICE

ED020720

ANOTHER LOOK AT TEACHER EVALUATION

IN

AMERICAN JUNIOR COLLEGES

by

Gordon Kilpatrick

March 1967

University of California, Los Angeles

Los Angeles, California

"The issue is not whether teachers will or will not be evaluated, but how the evaluation should be made and what uses should be made of the evaluation."[1]

--- A. S. Barr

WHY EVALUATE?

The purposes of evaluation are severalfold. Perhaps chief among any list of purposes which might be assembled would be communication--an otherwise cloistered administration finding out as accurately as possible what the school really is doing, how it is progressing both effort-wise and achievement-wise in pursuing its announced goals, and what skills are available to aid in pursuing its goals. Hardly of secondary import is the purpose of improving instruction. The process of improving instruction through evaluation would take the form of discovering and reinforcing teaching habits supportive of organizational goals and suggesting behavioral changes which will eliminate these habits which deter progress toward the achievement of these same organizational goals.[2] Communication of the good teaching habits also serves to improve instruction.

Another purpose of evaluation would serve the individual being evaluated. By periodically assessing the skills, attitudes, and knowledge of its teaching staff, the administration can encourage, promote,

---

[1]A. S. Barr (ed.), "An Expert Answers Questions Concerning Teacher Evaluation," Illinois Education, 49:377, May, 1961.

[2]Claude W. Fawcett, School Personnel Administration (New York: The McMillan Company, 1964), p. 61ff.

and assure the continued growth and development of the individual teachers in all of the above mentioned areas.

If faculties were made well aware of these primary purposes for evaluation, perhaps, they would be more amenable toward evaluation than is presently the case. As it is, mere talk of a proposed evaluation frequently provokes bristling resistance and sometimes anger and consternation on the part of some teachers.[3] What do they fear? If they are harboring inept teaching practices within the private sanctity of the classroom, they should welcome an opportunity to improve by discovering, through the evaluative process, exactly what needs to be changed. Such improvement may be impossible without the objective views and comments rendered by an evaluator, especially if the teacher is not proficient at self-evaluation. The haunting spectre of termination too often impairs an objective viewpoint towards evaluation by the teacher to be evaluated.

Disgruntlement is most often widespread particularly among the older members of the faculty, most of whom had tenure.[4] Could it be that they feared discovery of senility or lack of competence under the "umbrella of tenure"? If, indeed, Peterson was correct in proposing that evaluation should be a form of recognition of superior teachers,[5] the

---

[3]David G. Ryans, "Notes on the Rating of Teacher Performance," Journal of Educational Research, 47:9, May, 1954.

[4]David V. Tiedeman, "Teacher Competence and Its Relation to Salary," New England School Development Council, p. 37, July, 1956.

[5]Carl H. Peterson, "Seven Keys to Evaluating Teacher Competence," American School Board Journal, 136:34, May, 1958.

veteran faculty members might then take pride that their methods of instruction are being studied as a model so as to establish a more effective pattern of instruction for beginning faculty members to emulate. Evaluation can improve instruction by the process of sharing innovative ideas, procedures, or techniques. The evaluator himself, if he, too, is a teacher, may discover a new approach to an established concept. It may have never occurred to the evaluated teacher that his innovation was worth sharing. The evaluation thus offers a sounding board for expression of new techniques of teaching. Consequently, to maintain the spirit of the evaluation, a follow-up memorandum might be written communicating worthy impressions throughout the departmental faculty. Conversely, the evaluation may uncover fallacies, inadequate subject preparation, or improper procedures. In such cases, the evaluated teacher must be informed of his deficiency--preferably via a personal interview during which remedial action is suggested. After a suitable interval, a follow-up evaluation should be performed. Failure to correct deficiencies after an adequate time would require another personal interview with the teacher to find cause. Perhaps the sheer preponderance of behavioral changes required is so great as to obviate their attainment. If a workable plan of progressive attainment of remedial behavioral changes over a reasonable period of time is impossible while maintaining at least a minimal level of teaching adequacy, termination or reassignment are the only alternatives. The evaluated teacher should be given benefit of any doubt if the evaluation leads in this direction, especially if he is tenured. Older teachers, in whom

senility has reached a stage which deprives students of adequate instruction, may be diverted to more fruitful duties, outside of the classroom until they retire.

Keeping the above purposes for evaluation in mind, consideration is now given to two types of evaluation which differ in their function. First to be considered is the initial evaluation of non-employees as a part of the selection process. Attention will then be turned to evaluation of effectiveness of instruction of those teachers already employed.

Initial Evaluation Prior to Selection

In many junior colleges, teacher evaluation ends or receives only minor consideration after hiring. Since the initial evaluation is usually performed by an official at the administrative level (Dean of Instruction or higher), subsequent evaluations at lower levels tend not to differ with his judgment. The administrative evaluator may be so confident and self-satisfied with his own evaluation that he may consider subsequent evaluations unnecessary. Any combination of these situations is indeed unfortunate, especially if the administrative evaluator does not possess the degree of evaluative competence presumed, and in view of the fact that teachers, like anyone else, change over time. On the other hand, if the teacher does not change and the subject matter he teaches does, through evolution, then the teacher's knowledge is obsolete which too is meat for evaluation.

Consideration of current methodology in the initial evaluation is warranted at this point. After initial contact, the applicant sub-

mits transcripts, letters of recommendation or placement file, and a completed application. If perusal of these documents seem promising, the candidate is invited in for a personal interview. Based upon these sources of information and these alone, the applicant is evaluated and hired or rejected. The question might now well be, "Is this enough information to reach a valid decision on such a critical matter and is the information given proper consideration in reaching an evaluation?" To answer this, attention will be given to each of the sources of information: transcripts, letters of recommendation or placement file, completed application, and the interview.

Transcripts are an excellent measure of the applicant's ability, yet the evaluator too often submits them to only a superficial inspection, namely grade point average and number of hours in the field to be taught. As standards vary from school to school, so do grades. Sometimes these variances occur within a school depending upon the department. Thus, consideration should be given to these factors when evaluating. A direct correlation between grade point averages and intelligence or ability has not been proved conclusively, tending to detract somewhat from the high value that otherwise could be placed (and more often is assumed by most evaluators) on transcripts as a source of information for an accurate evaluation. Perhaps this results from an inability to obtain an accurate measurement of intelligence. On the other hand, a preliminary report of a study reported by Blake College claims that grade point averages increased with intelligence level only to a

point, and then began to decrease.[6] If this is so, a "B" student might be more intelligent than an "A" student, but was unchallenged and bored and hence did not use the full extent of his faculties. Also, what of the highly intelligent student whose ability to concentrate during college was hampered by a necessary job, or trauma, or other distractions particular to college-age youth? It should already be apparent to the reader that the usual amount of superficial scrutiny of transcripts is inadequate for an accurate appraisal of the applicant's ability. A deeper examination of transcripts in light of the applicant's experiences during the grading period would be appropriate. Scores on standardized tests will give indications of ability without the slant imparted by a particular institution's (or department's) philosophy of grading.

Letters of recommendation are useful in evaluating provided certain reservations are kept in mind. It must be remembered that the authors of the letters were selected by the applicant himself, and that it is only naturalthat he list those who will give favorable recommendations. Placement files, on the other hand, usually insist upon previous employers or former professors of the applicant as sources of recommendation; hence these are likely to be more objective in their appraisal and a more valuable source of information. Hopefully, a statement disclosing the quality of work performed will be included. If the evaluator is personally acquainted with any of those listed, a phone call may reveal insights not readily apparent from the written recommendation.

---

[6]Raymond Peat, "Blake --- A College Where Intelligence Matters," Interim, November, 1965, p. 3.

The completed application yields not only detailed personal information, but also a chronological record of previous experience. Job titles and summaries of duties reveal somewhat the degree of versatility or specialization of the applicant. This must be weighed with the employer's recommendation already considered. Stability is indicated by length of time spent on various jobs; "reason for leaving" is a point worthy of close consideration. (The evaluator would do well to consider construction of a new application form if the current form in use neglects these preceding items of information.) The applicability of his non-teaching experience to his usefulness as a teacher should be considered.

The interview is considered by some to be the least fruitful source of information.[7] However, in the quest for as much information as possible concerning the applicant's skills, attitudes, and knowledge, the truly conscientious evaluator leaves no potential source of information untried. The presentation proposed at conclusion of this paper will yield a wealth of information regarding teaching skills and extent of knowledge as well as other items not available elsewhere, but the interview is unexcelled as an instrument of assessing attitudes. If the applicant is made at ease and permitted to talk freely with occasional direction being given by pertinent questions, much information not available anywhere else concerning his attitudes is brought forth. A suggested set of questions which leads the interview into desirable

---

[7]Tiedeman, op. cit., p. 95.

informational areas is presented by Fawcett.[8] In case the presentation proposed is not used, the interviewer might consider inviting the applicant's potential immediate superior to share the interview in order to assess what he can relating to skills and knowledge. This would be especially desirable if the applicant's field were quite different from that of the interviewer.

Even if the preceding steps are followed carefully, the applicant's ability to communicate his subject matter effectively still has not been determined. Ability to present material in a logical order, on a level that is compatible with the ability of his students to understand and assimilate, and efficient utilization of available time are essential characteristics so far undetermined.

Effectiveness of Instruction

If evaluation of effectiveness of instruction could be accomplished, the junior college would be in a better position to know if its teachers are properly fulfilling their respective roles in the educational process. Such an evaluation could best be accomplished by observing the performance of students in their further endeavors and correlating this with the grade received in the course of the teacher being evaluated.

Evaluation of the protracted effectiveness of instruction is probably the most difficult to measure, but will probably be the most valuable piece of information once acquired. The difficulty stems from the sometimes scattered nature of the informational sources and the fact

---

[8]Fawcett, op. cit., pp. 33-36.

that the evaluation takes place significantly later than the actual instruction took place. This evaluation would be most easily accomplished in the transfer area of the curricula, since subsequent transcripts and professorial comments would be available. More difficult, but still not inaccessible, would be follow-up information on terminal students. Sources would be the solicited comments of employers or journeymen to whom graduates are apprenticed. Even the degree of success realized by the few terminating students who are self-employed would be useful information to the evaluator. Measuring effectiveness of instruction by this method would be most difficult if not impossible in the general education area, but since we are primarily concerned with teacher evaluation, and since junior college teachers seldom teach only one course, other avenues of evaluation are available to us. Compilation over the years of the correlations between the student's realized success and the teacher's prediction of success (the grade he gave the student) would say much of the teacher's effectiveness as an instructor; yet this avenue is largely ignored for teacher evaluation purposes although explored often on an institution-wide basis. The performance of the teacher's students on a standardized test in the particular subject area is another relatively unexploited avenue of teacher evaluation. Again, correlation with the students' assigned grades is required before a comparison with national norms is meaningful. Outside administration of these tests is desirable to prevent "prepping."

As professed earlier, instructional competence cannot be measured merely by assessing extent of knowledge of subject matter from

transcripts,[9] but results from the successful marriage of knowledge of subject matter and teaching skills, both of which are essential and neither of which is of value to any professed teacher without the other.[10] The successful teacher, steeped in his own knowledge, recognizes the level and capabilities of his students and organizes and presents the subject matter in a form and rate that is comprehensible to them. Thus, two sections of the same course may receive different presentations. Tailoring the presentation to the personality and readiness of the class is a sign of a successful teacher. Evaluating this facet of the teachers competence may be conducted in a classroom visit. The evaluator should place himself in the role of the student, assuming no greater knowledge than the average student at this stage of the course, to see if the material being presented is understandable. The faces of students in attendance offer an excellent indication of their degree of comprehension. The evaluator would thus be well advised not to sit too far in the rear, thus making this source of information unavailable to him. Students' comments or questions, if appropriate to the method of presentation, are also indicative of their comprehension and understanding and should not escape the attention of the evaluator. It is recognized that an artificial situation is created immediately upon the evaluator's entry into the classroom. Even if the teacher's attempt at atypical "best-foot-forward" type of behavior is successful or unsuccessful, ingrained

---

[9] Ryans, loc. cit.

[10] Adam M. Drayer, "Students' Views of the Qualifications of Their Teachers," Journal of Teacher Education, 12:3, September, 1961.

teaching habits are difficult to conceal.[11] The evaluator's alertness to detect these is required, especially if events leading up to a "critical incident" have been brought to his attention prior to the classroom visit. It may well be that the occasion of the evaluator's classroom visit is a phase of the exercise of the "critical incident technique," to be discussed more fully in the next section.

WHEN TO EVALUATE?

As already mentioned, evaluation is performed on applicants prior to selection. At what other times is evaluation timely?

Three occasions warrant teacher evaluation: (1) promotion consideration, (2) before awarding tenure, and (3) a "critical incident." These three occasions are not as distinct as they may seem; a critical incident may provoke the promotion consideration; awardance of tenure is considered a promotion in some systems.

Promotion Consideration

Promotions are meted out by various schools in various ways. The most widespread method is financial renumeration. A raise in pay above the customary annual increment may be justifiable awarded to an outstanding teacher. Since bonuses are not the rule in public schools and are expressly forbidden by law in most places, how else can excellence of instruction be rewarded monetarily? Some school districts in

---

[11] N. L. Gage, "The Appraisal of College Teaching," Journal of Higher Education, 32:1, January, 1961.

California award an extra annual salary increment to outstanding teachers. Any such reward system would provide incentive for all teachers to improve the effectiveness of their instruction leading to an overall improvement of instruction throughout the system. This should be reason enough to justify a small percentage of the budget to be designated as instructional incentive pay for this purpose. Objective standards must be used by the various evaluators in selection to assure that those teachers receiving this perquisite are only those who justly deserve it. To maintain the distinctiveness of the award, the relative percentage of teachers receiving it must be kept small. Only through thorough and objective teacher evaluation can the recipients of these awards be selected without partiality. On the other hand, if incompetence is detected and substantiated, the annual increment may be withheld in some districts, even if the teacher involved has tenure.

Merit pay has been discussed pro and con at length. The proposed correlation between teacher competence and salary was investigated and discussed at length by the New England School Development Council,[12] Taylor discussed prevailing attitudes for and against it.[13] An interesting merit pay system has been developed at Miami-Dade Junior College in Miami, Florida. The salary schedule has been discarded in theory in favor of salary ranges for each of the academic ranks. In practice,

---

[12]Tiedeman, loc. cit.

[13]James H. Taylor, "Whatever Became of Merit Rating?" The Personnel Man and His Job (New York: American Management Association, 1962), pp. 280-291.

the salary schedule still exists to the extent that teachers can predict exactly their pay since annual increments and promotion increments still follow the salary schedule guidelines faithfully. Theoretically, an outstanding teacher may receive a larger increment or earlier promotion than his colleagues on occasion when so recommended after evaluation by his department head. The administration in practice usually heeds time-in-rank requirements rather rigidly and such recognition of outstanding performance in the classroom by teachers is not general knowledge among the faculty.

A second method of promotion is via a system of academic rank. Although criteria for the various ranks are usually published, a formal evaluation for promotion is seldom performed. Unfortunately, the promotion date is usually determined by the calendar rather than by teaching proficiency, the time-in-rank requirements prevailing. Such requirements stifle the development of capable faculty members whose potential allows a faster rate of development than can be recognized by the usual time-in-rank requirements. Regular evaluations coupled with a promotion policy which rewards teaching proficiency without reference to the calendar would provide incentive for and recognition of the uninhibited development of a professorial talent.

Frequently promotion is made in the form of a department chairmanship or deanship. A different type of evaluation from that previously proposed is required since administrative ability rather than teaching ability is being evaluated. No claim of correlation between these

two factors is made. Exploration of this topic is beyond the scope of this paper.

In order not to penalize a deserving teacher by an organizational change just prior to his anticipated promotion, a superior who changes his duties should be required to conduct an evaluation of his immediate subordinates as one of his outgoing duties. The results should be recorded in writing, of course, preferably in a cumulative evaluation file. Thus the promotion date is not prolonged because of the incoming superior's reluctance to submit a promotion recommendation based on a short acquaintance or single evaluation.

Before Awarding Tenure

A thorough evaluation should be performed before awarding tenure.[14] Presumably, this will not be the first evaluation since hiring if the cumulative file proposed in the next subsection has been established and kept current. Tenure is usually awarded at the end of the third complete academic year of full-time service, but may be withheld for cause. It would behoove the administration to assure the degree of competence of a faculty member before awarding something as permanent and as far-reaching as tenure. An overhaul of the currently existing tenure policy to extend the probation period would greatly increase the reliability of the evaluation by allowing more time for more evaluation before awarding tenure.

---

[14]Barr, loc. cit.

The Critical Incident Technique

The critical incident technique was developed for use in industry by John C. Flanagan.[15] It has been defined for use in the educational field by Tiedeman as consisting of "the collection and analysis of reports of actual incidents in which the behavior of a teacher was judged by the reporter to be outstandingly effective or ineffective."[16] Ryans definition is essentially the same except that his reporter merely reports without judging and the incident must occur "in some specified teaching situation."[17] Corbally indicates that the critical incident approach is the most promising of any evaluation method, but recognized the shortcoming of inadequate means of measuring the degree of "criticalness" of the incident.[18] Flanagan identifies "critical" as it applies hereby saying, "To be critical, the incident must occur in a situation where the purpose or intent is clear to the observer and the consequences are sufficiently defined leaving little doubt concerning the effects."[19] How does the evaluator become aware of the occurrence of a critical incident, whether favorable or unfavorable? Must a network of spies be

---

[15]John C. Flanagan, "The Critical Incident Technique," Psychological Bulletin, 51:4, July, 1954.

[16]Tiedeman, op. cit., p. 61.

[17]David G. Ryans, Characteristics of Teachers (Washington: American Council on Education, 1960), p. 79.

[18]John E. Corbally, Jr., "The Critical Incident Technique and Educational Research," Educational Research Bulletin, 35:3, March 14, 1956.

[19]Flanagan, loc. cit.

established to furnish this information? By no means. The normal operation of the school or junior college ordinarily provides this information through both formal and informal communication channels. Sources may include the evaluated teacher himself, other teachers, supervisors, student teachers, education students, deans, and principals or presidents. Some of these may be appointed evaluators reporting the results of formal classroom evaluation. Students, too, are a source of information, but students who volunteer such information generally are self-selected complainers. The reliability of such information is open to question, but significant numbers of similar reports establishes a trend which lends credence to itself. Although a critical incident may be reflected in the results of student opinion questionnaires (discussed later in this paper), a critical incident is a particular occurrence which should not be confused with the generalized student opinion. Student opinion questionnaires are therefore not reporting sources of critical incidents. The distinguishing features of the critical incident techniques is that each item of information concerns a specific evaluator. Flanagan recommends that periodic, even daily written entries be made on the performance record, claiming that half of the details of the incidents are forgotten in one week and three-fourths are forgotten in two weeks.[20] Woodburne, too, recommends the creation and maintenance of such a cumulative evalua-

---

[20]John C. Flanagan, "Personnel Research and the Better Use of Human Resources," The Personnel Man and His Job (New York: American Management Association, 1962) pp. 348-349.

tion file, although he does not relate it to critical incidents.[21] The source of each written entry should be recorded accurately so as to indicate the reliability which may be attached to the entry. Hearsay, without recognizable source, should not be recorded.

Question of Competence

Sometimes a question of professorial competence is raised, either by the general public or by one or more students. The cumulative file of critical incidents and evaluation results stretching over a period of time up to the most recent (hopefully within the current school term) should be available for the private reference of the administrator hearing the complaint so that he is better able to judge the merits of the case. Thus the administrator is able to speak promptly and knowledgeably to the public of the degree of competence of the faculty as a whole or of one faculty member in particular. If necessary, a supplemental evaluation may be performed to update the file. Lack of such information could easily lead to an embarrassing situation.

Flanagan suggests the following steps in the critical incident technique:

1. Determination of the general aim of activities
2. Development of plans and specifications for collecting factual incidents
3. Collection of data
4. Analysis of data
5. Interpretation and reporting.

Flanagan indicates that the evaluator does not have to necessarily be

---

[21] L. S. Woodburne, "Suggest Criteria for Academic Promotion," American Association of University Professors Bulletin, 29:2, April, 1943.

the observer, but that observers must be carefully trained to report facts, not judgments. Thus, the observation step, step (3), is an objective one. It must be pointed out that the observer does indeed use his judgment in selecting WHAT to report; he is in effect an abstracter.[22] Step (4) is less objective, and step (5) subjectively requires value judgments, but require accurate reporting of the biases used in arriving at the final interpretation.[23] This interpretaion is NOT in itself an evaluation, nor does it reveal "successful" teaching; it merely describes teaching. The critical incident technique is therefore the rationale for evaluation. The evaluator must supply the criteria for evaluation. He must, therefore, possess a second order set of criteria to evaluate the values to be used. Without this, he is seeking to recognize something he does not know.[24]

Fawcett's method of using the critical incident technique to provoke evaluation is illustrative of the above points. He contends that the time of evaluation is indicated by "signals," which are, in effect, critical incidents. Evaluation should take place as soon as possible after the critical behavior has been displayed by the teachers. The aspect to be studied in the evaluation may be vaguely suggested by the critical incidents. Via a conference, both the evaluator and the

---

[22]Donald M. Medley, "Measuring Classroom Behavior by Systematic Observation," Handbook of Research on Teaching (Chicago: Rand McNally Co., 1967), p. 257.

[23]Flanagan, loc. cit.

[24]Hobart W. Burns, "Success Criteria and the Critical Incident Technique," Phi Delta Kappan, 38:73-75, November, 1956.

teacher to be evaluated seeks to develop a probable hypothesis to be investigated. This is a rather subjective phase of the evaluation. After formulation, the hypothesis is tested by observation. The observation is conducted very objectively. If the hypothesis does not withstand the test, an alternative is mutually formulated and the process repeated. A hypothesis which holds is then subject of the next evaluator-teacher conference at which a course of action is decided upon to either commend or redirect the evaluated's behavior. Required remedial action may well be the basis for in-service training project.[25]

Prior to evaluation or establishment of an in-service training project must be a clear definition of expectations of teacher behavior. Not only is it impossible for an evaluator to report that a person is effective or ineffective by performing specific activities without the evaluator knowing what is expected of the teacher,[26] but it is also impossible for the teacher to operate effectively in this vacuum.[27] It is presumptuous to assume that teachers know exactly what is expected of them, such information having been derived entirely from their job titles, just as it would be presumptuous to assume all teachers teach with optimum effectiveness (which would obviate evaluation entirely). Even if they did, it would not be assured that the expectations of teacher behavior as derived from job titles by the evaluator would be the same.

---

[25]Fawcett, op. cit., p. 57.

[26]John C. Flanagan, "The Critical Incident Technique," Psychological Bulletin, 51:4, July, 1954.

[27]Medley, loc. cit.

Therefore, it seems logical that a written record of expected behavior and standards, a job description, should be prepared for each position and disseminated to all concerned. Vagueness and generalities should be avoided in preparing job descriptions. Specific acts which contribute to effectiveness and/or success should be delineated as clearly as possible. Such would serve not only to guide teacher behavior, but also to give direction to in-service training projects.

The interesting point in the timing of evaluation by means of the critical incident technique would be truly mediocre teachers, neither good nor bad, who seldom encounters an incident which may be labelled "critical." To say there are always critical incidents is not enough, for if the incidents are so widely spaced so as never to constitute a "signal," when is the teacher to be evaluated? Supplementing the critical incident method of indicating need for evaluation should be a regulation which requires periodic evaluations to fill time gaps not covered by the aforementioned procedure. All teachers are capable of improvement most especially mediocre teachers. To deny a mediocre teacher an opportunity to discover and remedy his shortcomings and thus improve instruction, merely by not evaluating is as neglectful as ignoring inferior instruction by the same mechanism.

The evaluated teacher may feel that the evaluator or whoever is prying if he enters the former's classroom without being required to do so;[28] the evaluator or observer knows this which contributes to

---

[28] Medley, op. cit., p. 247.

some degree of embarrassment and a strained interpersonal relationship. Regulations would ease this situation. Performance of general evaluators (not the critical incident approach) without classroom observations contributes to laxity, the teacher being judged perhaps on his demeanor in his office, his affability in the coffee klatch, or his participation and performance on various academic committees. Indeed, all of these are important to the make-up of the whole teacher, and should be considered in an over-all evaluation of that teacher.

WHO SHOULD EVALUATE AND HOW?

In many junior colleges evaluation is made by the administration. Reason for this lies in the organizational structure wherein the instructional staff is answerable directly to the Dean of Instruction without intervening department heads. Unfortunately, such evaluations are usually based upon transcripts, interviews, and other applicant information. If a classroom visitation is made by the Dean of Instruction, what if the course being presented is outside of his subject field? His very inability to understand a presentation in a course such an integral calculus (through inadequacy of his own background in higher mathematics) may lead him to downgrade the teacher in the evaluation, even though the presentation is entirely understandable by the students. Even though he may be able to assess teaching skills and attitudes, could not this too be assessed by an evaluator competent to assess knowledge? Individual classroom visitation of a large faculty

would demand large amounts of the Dean's time, due to this unfortunate organization. If this situation is unworkable, who then should evaluate?

Preferably the evaluator should be as nearly as possible of the same background as the teacher being evaluated. Such people would most likely be found within the same department. Intradepartmental colleagues would have such similar backgrounds. Some schools have evaluation committees set up within departments for this very purpose. However several problems develop. Personal friendships or petty rivalry and jealousy may deter an objective evaluation. Colleagues may even tend to judge each other on "coffee-klatch behavior." Also, who would evaluate the evaluators? Reciprocal evaluations would tend to foster alliances or the converse which could hardly yield truly objective evaluations.

The Army's method of evaluations being performed by senior officers is a valid one. The one person of the same background but still senior to the evaluated teacher is the <u>department head</u>. He is in a position to view all aspects of professional development and behavior. He sees the evaluated in the classroom, in the office situation, in the coffee klatch or lunch room, at professional meetings, and possibly socially on occasion. Thus he is able to draw together many observations of various aspects of the evaluated teacher's state of professional development. Being of the same background, he is very able to judge the teacher's organization and appropriateness of presented materials and to assess the currentness of his own knowledge. The latter is especially important in the rapidly evolving scientific fields.

If, as in most junior colleges, he is in effect a teaching administrator, his intermediate position gives him a viewpoint vastly superior for evaluation purposes than the viewpoint of any other available staff member. In case evaluation is performed by a committee, he is an indispensable member and his judgment should carry much weight.

But who evaluates the department head?

Since the department head himself is usually a teaching-administrator, his evaluation should be of an administrative type in addition to an evaluation in his teaching area. This evaluation is best performed by his next higher superior who has day-to-day contact with him in administrative areas. Likely this is a dean or division director who has adequate knowledge at his disposal to render an evaluation without involved investigative procedures. This ease of evaluation should not contribute to laxity; the evaluation should be performed and recorded in writing periodically.

Much has been said for and against sampling students' opinions of faculty members for the purpose of evaluation. Student rating of faculty finds widespread acceptance on campuses of four-year colleges and universities; four out of five colleges which tried student rating of faculty (N=296) retained the system.[29] Yet, at the junior college level especially, the question of the students' competency to judge faculty has been raised again and again. The reader might well recall at this point

---

[29]F. J. Mueller, "Trends in Student Ratings of Faculty," American Association of University Professors Bulletin, 37:2, Summer, 1957.

his own evaluation of various professors when he himself was a lower division student; how valid were these evaluations? It is Williams' opinion that students are an excellent source of evaluative information,[30] After all, they see more of the instructor in the actual teaching situation than anyone else and are thus in the best position to assess the effectiveness of instruction.[31] They have observed the typical behavior of the instructor in typical teaching situations and are in a vastly superior observational position when compared to the classroom visitor who views an artificial, atypical teaching situation. They know, better than anyone else, if the instructor was adept at presenting information in an understandable way. McKeachie sums it up by saying:

> Regardless of the validity of the student ratings, however they may indicate the students' credence in the instructor's statements and thus his effectiveness in bringing about attitude changes.[32]

Naturally the student is incapable of rendering a full evaluation phrased in the professional jargon of the educator. Even those who are not skillful in expression have opinions which may be expressed more objectively. Hence it is not too surprising that most student opinion questionnaires are objective in type. Usually they are distributed to

---

[30]Glenn D. Williams, "Your Students Can Help You Be a Better Teacher," Illinois Education, 50:284, March, 1962.

[31]Priscilla R. Meyer and R. M. Patton, "Can Student Rating of Instructors Be Painless and Foolproof?" School and Society, 80:200-201, December 25, 1954.

[32]W. J. McKeachie, "Research on Teachers at the College and University Level," Handbook of Research on Teaching (Chicago: Rand McNally & Co., 1963), p. 1130.

the students toward the end of the course. The anonymity of the students must be preserved; the students must be fully informed and convinced that their responses to the questionnaires will not in any way affect their grade in spite of the fact that several studies had shown no correlation between the students' ratings and grade expected.[33] Thus it might be well to have some person other than the evaluated teacher collect the completed questionnaires which are then delivered directly to the evaluator. The evaluated teacher should not have access to the contents until after final course grades for all members in the class have been assigned and recorded.[34] This should be known to the students.

The manner in which information derived from student questionnaires is used may be varied. It may be analyzed and used by the department head in arriving at his overall evaluation, or it may be used along with the department head's independent evaluation by the dean in his overall evaluation. The weight given to student opinion may vary also, as long as it is consistent within the school for a particular evaluation period. Thus even those who consider student ratings inept would know that some flexibility is afforded to the evaluator who with his presumed

---

[33]Alexis M. Anikeeff, "Factors Affecting Students' Evaluation of College Faculty Members," Journal of Applied Psychology, 37:6, December, 1953; Earl Hudelson, "The Validity of Student Rating of Instructors," School and Society, 73:1897, April 18, 1951; V. N. Voeks and G. M. French, "Are Student Ratings of Teachers Affected by Grades?" Journal of Higher Education, 31:6, June, 1960.

[34]Carl H. Weaver, "Instructor Rating by College Students," Journal of Education Psychology, 51:1, 1960.

greater wisdom may judge the extent to which student ratings should be used in the total evaluation.

The construction of the questionnaire should be carefully considered. Since subjective qualities are being weighed, a method of adding objectivity to the responses to make it more processable would be desired. This will enable averaging of data so as to obtain a semi-quantitative rating of the teacher. Presumably, the questionnaires are screened before processing to eliminate the unconscientious questionnaires, those that denote consistent extreme views on all responses, be they good or bad. Hopefully, the number of this type will be very small. A significant number of one kind with an inclination established in the same direction by the remainder should not discount the students' opinion, but should signal the evaluator that additional attention is warranted in his evaluation of the teacher in question. This would constitute a critical incident. Some teachers experience popularity (or lack thereof) among the students due to the grading patterns they have established. The teacher who is generous with high grades enjoys a popularity not experienced by his counterpart who gives A's and B's sparingly, or even by the just teacher who gives approximately equal numbers of both high and low grades. It was found by Weaver that this "popularity halo" did not affect the ratings students gave their teachers, that students indeed rated teaching skills and abilities, not the teacher's personality.[35] On the other hand, the

---

[35]Weaver, loc. cit.

students personality did affect the way he rated his teacher.[36] Cynamon found students' responses varied according to the announced intended use of the questionnaire.[37] Gage found a correlation between rating given and level of the class, whether the class met on or off campus, the size of the class, and whether it was elective or not.[38]

An an adjunct to student opinion determination may be alumni opinion determination. Likewise a constructed questionnaire would be required similar to the one just described. The difficulty herein lies in the problem of dissemination. Many students move after graduation. Some are apathetic or laggard in responding. One hundred per cent (100%) response to a mailed questionnaire is virtually unknown. The follow-up study mentioned earlier in this paper should prove a more lucrative source of reliable information. Drucker found that alumni opinion of the instructor varied little from student opinion;[39] therefore since students are much more readily available, the mechanics of the problem would favor using student opinion rather than alumni opinion.

---

[36]B. N. Phillips, "Authoritarian, Hostile, and Anxious Students" Ratings of an Instructor," California Journal of Educational Research, 9:1, January, 1960.

[37]Manuel Cynamon and Shirley U. Wedeen, "Emotional Factors in the Reliability of Student Rating of Teachers," Journal of Educational Research, 51:8, April, 1958.

[38]Gage, loc. cit.

[39]A. J. Drucker and H. H. Remmers, "Do Alumni and Students Differ in Their Attitudes Towards Instructors?" Journal of Educational Psychology, 42:3, March, 1951.

After perusal of many student rating questionnaires,[40] a more workable questionnaire was formulated by this author and is found in the appendix. The "forced-choice" type of questionnaire advocated by many[41] was considered but not used. It was thought that this type bounded the extremes by choice of descriptive phrases used by the questionnaire writer. Rather, the student is able to decide in his own mind what constitutes an extreme as he evaluates using a letter grade system on the teacher not unlike that used by the teacher on the student. The proposed questionnaire is readily adaptable to computer tabulation so as to cut down processing time and clerical expenditure. It is divided into sections so that inadequate areas (in the students' opinions) may be assigned to the different areas. The number of items is kept small so that the questionnaire may be administered within a reasonably

---

[40]Edward J. Durnall, Jr., "The Student Evaluates the Teacher," Junior College Journal, 27:5, January, 1957; Lynn H. Harris, "A Device for Student Evaluation of a Course," Junior College Journal, 16:1, September, 1945: Robert L. Isaacson et al., "Correlation of Teacher Personality Variables and Student Ratings," Journal of Educational Psychology, 55:6, December, 1964: H. H. Remmers, "Rating Methods in Research and Teaching," Handbook of Research on Teaching (Chicago: Rand McNally & Co., 1967), pp. 329-378; Franz Schneider, "More Than an Academic Question," The Pestalozzi Press, Berkeley, California, 1945; Williams, loc. cit.

[41]G. D. Lovell and C. F. Haner, "Forced-Choice Applied to College Faculty Rating," Educational and Psychological Measurement, 15:3, 1955; Rollin N. Patton and P. R. Meyer, "A Forced-Choice Rating Form for College Teachers," Journal of Educational Psychology, 42:3, March, 1951; Ryans, loc. cit.

short time. The descriptive terms are kept concise so that the student may spend more of his time considering his responses rather than most of his time reading.

The proposed questionnaire is not considered the final answer. It is subject to refinement and to adaption to varying schools or departments. It is merely a suggestion. Even if a school-wide evaluation program is not in effect, the instructor may use this form in his own classes for self-improvement by detecting and eliminating any stumbling blocks to effective instruction.

Many such stumbling blocks have been detected via student rating sheets. The most common one was belittling the students.[42] This particular trait would most likely be guarded against and not be detected by an evaluator on a classroom visit. More subtle faults detected by student ratings were: dwelling upon the obvious, repeating too much, voice inadequacies, unfair grading, and unfair tests. Even among the top rated teachers, faults were found, most commonly speaking too fast and assuming too much background.[43] Students may not be completely aware of what constitutes a good teacher, but perhaps they are better able to detect a bad teacher.[44]

---

[42] Voeks, loc. cit.

[43] Voeks, loc. cit.

[44] Drayer, loc. cit.

RECOMMENDED GUIDELINES FOR TEACHER EVALUATION

Initial Evaluation: (These are self-evident in the text of this paper and will not be repeated here.)

Subsequent Evaluations:

1. General evaluations should be conducted periodically, prompted by critical incidents or required by regulation, and performed by the evaluator who should be the department chairman of the evaluated teacher. Classroom visitation should be included so that method of presentation, organization of materials, assessment of degree of preparation, and ability to hold interest of students may all be evaluated. A check list of these items and others deemed appropriate should be used. The teachers behavior will be atypical due to the evaluator's presence, but ingrained habits will be in evidence.
2. Specific evaluations performed according to the methods of Fawcett[45] should also be performed as required.
3. Student opinions should be sampled via a simple computer scored questionnaire. Faked answers by unconscientious students should be eliminated by preliminary screening before processing.
4. Committee work, publications, research activity, professional participation, and continued scholarship are self-evident and should be a part of the evaluation. Since junior colleges are

---

[45] Fawcett, loc. cit.

primarily teaching organizations, the weight of this part of the evaluation should be less than the first two mentioned.

5. All findings should be recorded in writing in an individual cumulative evaluation file which may be the same as the personnel file. It should be brought up to date at least quarterly during employment incorporating results of above investigations into one overall evaluation. Outstanding teachers should be rewarded. Deficient teachers should be told of the nature of their deficiency so that they may take remedial action before their next evaluation. Communication of improved methods of instruction observed during evaluation should be made to all of the instructional staff. Weighings of the various parts of the evaluation should be decided before the evaluation begins. The focal point of all parts of the evaluation may be at either the department level, division level, dean level, or presidency level depending upon the size and orientation of the junior college.

Presentation: A seminar program should be established departmentally so as to give each faculty member a periodic opportunity of presenting an interesting topic of his choice within his field to the rest of the department. Students should be invited for subject matter content. This would give all concerned with evaluation the opportunity of observing each faculty member in action within his field. Method of presentation, organization of material, presence of distracting mannerisms, and adequacy of voice could all be evaluated as well as the more apparent subject-matter recency and proficiency. The same

type of presentation should be used in final screening of applicants to determine the suitability and aid in selection.

EPILOG

Transcending this entire paper has been an underlying concept, that those characteristics which constitute a good instructor and distinguishes him from a bad one are known at least qualitatively by the evaluator. It has been stated that subject matter knowledge and teaching proficiency are primary, yet allied with these are the countless little details such as personality, integrity, dedication, interest, enthusiasm, etc., which constitute a teacher's attitude and which vary in their combination to either contribute to or deter effective instruction. Ultimately, it is the evaluator who decides whether an individual's particular combination of ingredients is beneficial or detrimental to effective instruction and to what extent. It has been the purpose of this paper to suggest guidelines and their rationale to assist the evaluator in reaching his decision on a somewhat semi-quantitative basis. To suggest criteria to be used in evaluation is beyond the scope of this paper.

(RECOMMENDED)

STUDENT OPINION QUESTIONNAIRE

This questionnaire is to determine your opinion of your teachers method of teaching so as to improve future instruction. Your response will in no way affect your grade in this course since the teacher will not see the results until after final grades are recorded. Fill in the blanks below but do not indicate your name anywhere on this paper.

Teacher's Name _____ Course Title _____

Section Number _____ Enrollment _____ Term _____, 19 ___

Indicate your opinion of the items below by letter.

A = excellent, B = above average, C = average, D = below average,

F = utterly failing.

If you wish to elaborate on any response, record the item number and your comments on a separate sheet of paper.

1. Teacher's mastery of subject.
2. Teacher's organization of material.
3. Understandability of presentation.
4. Relevance of material presented in class.
5. Relevance of outside assignments.
6. Teacher's awareness of students level of understanding.
7. Teacher's respect of students' knowledge.
8. Teacher's attitude toward students.
9. Teacher's enthusiasm toward subject.
10. Teacher's apparent degree of dedication.
11. Teacher's willingness to answer questions and explain.
12. Teacher's availability for individual help.
13. Teacher's patience, self-control, and temperament.
14. Opportunity for discussion.
15. Teacher's interest in class progress.
16. Teacher's ability to create interest in subject matter.
17. Teacher's ability to encourage students to think for themselves.
18. Teacher's ability to encourage students to be creative and try new ideas.

19. Teacher's ability to make students clearly aware of his expectations.
20. Teacher's abilility to use class time efficiently.
21. Distribution of tests.
22. Fairness of tests.
23. Fairness of grading.
24. Promptness of returning graded tests.
25. Distribution of work during term.
26. Teacher's appearance.

For the remaining questions, use the following response code:

A = excessive, B = more than average, C = about right,
D = less than average, F = inadequate.

27. Amount of total work required for this course.
28. Amount of outside study required.
29. Time required by this course.
30. Number of tests.
31. Length of tests.
32. Coverage of tests.
33. Depth of knowledge of subject matter required on tests.
34. Teacher follows textbook.
35. Integration of appropriate outside material.
36. Degree of formality in class.
37. Amount of repetition.
38. Scope of course as presented by this teacher.
39. Depth of course as presented by this teacher.
40. Using the former response code, A = excellent, B = above average, etc., what overall evaluation would you assign to this teacher?

BIBLIOGRAPHY

American Management Association, The Personnel Man and His Job, New York: American Management Association, 1962.

Anikeeff, Alexis M. "Factors Affecting Students' Evaluation of College Faculty Members," Journal of Applied Psychology, 37:6, December, 1953.

Burns, Hobart W. "Success Criteria and The Critical Incident Technique," Phi Delta Kappan, 38:73-75, November, 1956.

Corbally, John E., Jr. "The Critical Incident Technique and Educational Research," Educational Research Bulletin, 35:3, March 14, 1956.

Cynamon, Manuel and S. U. Wedeen. "Emotional Factors in the Reliability of Student Rating of Teachers," Journal of Educational Research, 51:8, April, 1958.

Drayer, Adam M. "Students' Views of the Qualifications of Their Teachers," Journal of Teacher Education, 12:3, September, 1961.

Drucker, A. J. and H. H. Remmers. "Do Alumni and Students Differ in Their Attitudes Towards Instructors," Journal of Educational Psychology, 42:3, March, 1951.

Durnall, Edward J., Jr. "The Student Evaluates the Teacher," Junior College Journal, 22:6, February, 1952.

_____(ed.), "An Expert (A. S. Barr) Answers Questions Concerning Teacher Evaluation," Illinois Education, 49:9, May, 1961.

Fawcett, Claude W. School Personnel Administration. New York: The McMillan Company, 1964.

Flanagan, John C. "The Critical Incident Technique," Psychological Bulletin, 51:4, July, 1954.

Flanagan, John C. "Personnel Research and the Better Use of Human Resources," The Personnel Man and His Job. New York: American Management Association, 1962.

Gage, N. L. "The Appraisal of College Teaching," Journal of Higher Education, 32:1, January, 1961.

_____(ed.), Handbook of Research on Teaching. Chicago: Rand McNally & Company, 1963.

Harris, Lynn H. "A Device for Student Evaluation of a Course," Junior College Journal, 16:1, September, 1945.

Hudelson, Earl. "The Validity of Student Rating of Instructors," School and Society, 73:1897, April 28, 1951.

Isaacson, Robert L. et al. "Correlation of Teacher Personality Variables and Student Ratings," Journal of Educational Psychology, 55:6, December, 1964.

Lovell, G. D. and C. F. Haner. "Forced-Choice Applied to College Faculty Rating," Educational and Psychological Measurement, 15:3, 1955.

McKeachie, A. J. "Research on Teachers at the College and University Level," Handbook of Research on Teaching. Chicago: Rand McNally & Company, 1963.

Medalie, Richard J. "Grading the Teachers," Phi Delta Kappan, 31:5, January, 1950.

Medley, Donald M. "Measuring Classroom Behavior by Systematic Observation," Handbook of Research on Teaching. Chicago: Rand McNally & Company, 1963.

Mueller, F. J. "Trends in Student Ratings of Faculty," American Association of University Professors Bulletin, 37:2, Summer, 1957.

National Education Association. Better Than Rating. Washington D.C., 1950.

Patton, Rollin M. and P. R. Meyer. "A Forced-Choice Rating Form for College Teachers," Journal of Educational Psychology, December, 1955.

Peat, Raymond. "Blake--A College Where Intelligence Matters," Interim, November, 1965.

Peterson, Carl H. "Seven Keys to Evaluating Teacher Competence," American School Board Journal, 136:5, May, 1958.

Phillips, B. N. "Authoritarian, Hostile, and Anxious Students' Ratings of an Instructor," California Journal of Educational Research, 9:1, January, 1960.

Rauch, Walter E. "Instructor Rating Sheets," Junior College Journal, 22:8, April, 1952.

Remmers, H. H. "Rating Methods in Research on Teaching," Handbook of Research on Teaching. Chicago: Rand McNally & Company, 1963.

Report of the Committee on College and University Teaching. American Association of University Professors, 1933.

Riley, John W. and B. F. Ryan. "The Student Looks at His Teacher," Rutgers University Press. New Brunswick, New Jersey, 1950.

Ryans, David G. Characteristics of Teachers. Washington: American Council on Education, 1960.

Ryans, David G. "Notes on the Rating of Teacher Performance," Journal of Educational Research, 47:9, May, 1954.

Schneider, Franz. "More Than an Academic Question," The Pestalozzi Press. Berkeley, California, 1945.

Taylor, James H. "Whatever Became of Merit Rating?" The Personnel Man and His Job. New York: American Management Association, 1962.

Taylor, W. S. "Student Ratings of Instruction," Journal of Higher Education. 18:4, April, 1947.

Tiedeman, David V. "Teacher Competence and Its Relation to Salary," New England School Development Council, July, 1956.

Voeks, V. W. and G. N. French. "Are Student Ratings of Teachers Affected by Grades?" Journal of Higher Education, 31:6, June, 1960.

Weaver, Carl H. "Instructor Rating by College Students," Journal of Educational Psychology, 51:1, 1960.

Williams, Glenn D. "Your Students Can Help You Be a Better Teacher," Illinois Education, 50:7, March, 1962.

Woodburne, L. S. "Suggested Criteria for Academic Promotion," American Association of University Professors Bulletin, 29:2, April, 1943.